Araştırma Makalesi
(Research Article)

Ege Üniv. Ziraat Fak. Derg., 2016, 53 (3):245-250
ISSN 1018 – 8851

Çiğdem TAKMA[1]
Öznur İŞÇİ GÜNERİ[2]
Yakut GEVREKÇİ[1]

Investigation of Stillbirth Rate Using Logistic Regression Analysis in Holstein Friesian Calves

Siyah Alaca Buzağılarda Ölü Doğum Oranının Lojistik Regresyon Analizi İle İncelenmesi

[1] Department of Animal Science, Faculty of Agriculture, University of Ege, 35100, Izmir /Turkey
[2] Department of Statistics, Faculty of Science, University of Muğla Sıtkı Koçman, 48300, Muğla, Turkey

corresponding author: cigdem.takma@ege.edu.tr.

Alınış (Received):12.11.2015 Kabul tarihi (Accepted): 08.04.2016

Key Words:

Holstein Friesian, stillbirth, logistic regression, calf

ABSTRACT

Logistic regression analysis is a method to determine the reason-result relationship of independent variable(s) with dependent variable, which has binary or multiple categorical structures. In this study, sex of calf, parity and calving year-season effects on stillbirth were analyzed with binary logistic regression analysis. Study material was obtained from the USA National Association of Animal Breeders collected among 2003-2005 with a total of 404460 birth records of single born calves. According to the results, sex of calf, parity and calving year-season effects on stillbirth were found statistically significant (P<0.05). The model showed good fit, based on Hosmer-Lemeshow goodness of fit statistics (P>0.12). When all variables were analyzed together in the same model, stillbirth rate of female calves compared to male calves was found to be more than 1.03 times higher. In addition, risk of stillbirth was decreased by increasing parity. On the other hand, the risk of stillbirth in summer calves was found to be higher than winter calves. In our country, data sets on stillbirth rates should be collected and risk factors that have an effect on stillbirth must be detected and then calf deaths could be controlled here, too.

Anahtar Sözcükler:

Siyah Alaca, ölü doğum, lojistik regresyon, buzağı

ÖZET

Lojistik regresyon analizi ikili veya çoklu kategorik yapıdaki bağımlı değişkenin bağımsız değişkenlerle neden sonuç ilişkisini belirlemede yararlanılan bir yöntemdir. Bu çalışmada buzağı ölü doğum oranları üzerine buzağı cinsiyeti, laktasyon sırası ve buzağılama yıl-mevsim etkileri ikili lojistik regresyon analizi ile incelenmiştir. Çalışmanın materyali ABD Ulusal Hayvan Yetiştiricileri Birliği'ne kayıtlı 2003-2005 yılları arasında yetiştirilen toplam 404460 adet tek doğan buzağının doğum kayıtlarıdır. Analiz sonuçlarına göre buzağı cinsiyeti, laktasyon sırası ve buzağılama yıl-mevsim değişkenlerinin buzağı ölü doğumlarına etkileri önemli (P<0.05) bulunmuştur. Hosmer-Lemeshow uyum iyiliği testi ile modelin iyi uyum gösterdiği belirlenmiştir (P>0.12). Tüm değişkenler aynı modelde birlikte analiz edildiğinde, dişilerdeki ölü doğum oranının erkeklere göre 1.03 kat daha fazla olduğu belirlenmiştir. Ayrıca, laktasyon sırasının artmasıyla ölü doğum oranının azaldığı saptanmıştır. Diğer yandan, yaz doğumlarındaki ölü doğum oranı kış doğumlarından fazla bulunmuştur. Ülkemizde ölü doğum oranına ilişkin kayıtlar toplanmalı ve bu oran üzerine etkili olabilecek risk faktörleri belirlenmeli ve böylece buzağı ölümleri kontrol altına alınabilmelidir.

INTRODUCTION

Several methods are used to explain the cause-result relationships in scientific research. Although the choice of method is significantly related with structure of the variables in the research, often simple or multiple regression analysis can be used. In the data set that was applied to these methods, dependent (Y), independent variables (X) and the error term must be normally distributed. On the other hand, sometimes the dependent variable takes two (binary - dichotomous) or more than two-class (polychotomous) whereas in fact it should be a continuous variable (Sharma, 1996). In this case the regression assumptions cannot be provided (Johnson and Wichern, 2005) and Least Squares parameter estimates lose the properties of best linear and unbiased estimator (BLUE) (Park, 2010).

In such cases logistic regression analysis is preferred. Because logistic regression analysis is a method with independent variables, even a combination of both of them is continuous or discrete and dependent variable is discrete (Antonogeorgos et al., 2009; Çokluk et al., 2010). Logistic regression analysis can be divided into three groups depending on the structure of the dependent variable. If structure of dependent variable is categorical with two groups: binary, if classified with more than two groups: nominal, if it has ranking scale: ordinal logistic regression analysis is used (Cook et al.,, 2001; Stephenson, 2008). Recently, logistic regression analysis is increasingly common in all disciplines. It's easy using the development of parameter estimation methods, the availability of software and interpreting the results in a meaningful way are the main causes of this interest.

There is not enough research in which stillbirth rates of calves are analyzed by logistic regression analysis. Yakubu et al., (2014) examined the breed, season, parity, and litter number on effects of abortion and stillbirth in 5,268 goats from four different goat breeds in Nigeria. As a result of binary logistic regression analysis, the related factors have been identified of the abortion and stillbirth risk factors. Zadeh (2014) studied the effect of dystocia on stillbirth from a total of 16 herds and 104,572 Holstein calf records with logistic regression analysis. According to the results of the study, in cows with birth difficulties the incidence of stillbirth increased, and yield characteristics decreased. Similarly, Atashi (2011) investigated the effects of risk factors on stillbirth and stillbirth effects on lactation performance in 5,201 Holsteins. According to logistic regression analysis results, calving year, parity and dystocia increases the incidence of stillbirth, whereas there is no effect by calving season. Also, stillbirth is high in the first calving and dystocia increases the risk of stillbirth. Meyer et al. (2001) investigated the mortality gestation period, calving year, calving season, sex, parity and dystocia effects on 666,341 Holstein's stillbirth rates of multiparous and primiparous cows in the USA. These factors were significant. Bicalho et al. (2007) revealed that sex of calf, parity and dystocia had significant effects on 13,608 Holstein calves reared in the UK.

In this study, the effects of calf sex, parity, calving year and season on stillbirth were examined according to the binary logistic regression analysis and are intended to remedy the lack of literature on this subject.

MATERIAL and METHOD

Material

The material of this study consists of Holstein calf birth records collected within the period 2003 to 2005 from 3,980 herds belonging to members of US National Animal Breeders Association in Columbia and Missouri. The calves of 404,460 Holstein cows with 15 different parities were scored as calf live (0) and stillbirth (1) if they were live or defined as death within 48 h from parturition, respectively.

The months of collected birth records were classified as May to September (1, summer), October to April (2, winter) and years grouped as 2003 (1) 2004 (2) and 2005 (3). These groups were combined into a single year-season categorical variable (2003/May-September-1, 2003-2004/October-April-2, 2004/May-September-3, 2004-2005/October-April-4 and 2005/May-September-5). Also, due to the lack of observations in the six and subsequent lactation records were evaluated under the sixth lactation.

Method

Logistic regression analysis uses maximum likelihood method rather than the least squares method. The values of the estimated parameters are adjusted iteratively until the maximum likelihood value for the estimated parameters is obtained. That is, maximum likelihood approaches try to find estimates of parameters that make the data actually observed most likely (Hair et al., 2006). Moreover, the probability of logistic regression analysis is based on the odds ratio and logarithm of the odds. Odds ratio is defined as the ratio of the probability that an event will occur divided by the probability that the event will not occur (Mertler and Vannatta, 2005). In other words, odds ratio regression represents the odds change in the dependent variable for a change of one unit in the independent variable and tells us how many times more likely the event is to happen than not happen.

This interpretation is similar to that of the linear regression coefficient. The only difference is that the change in the dependent variable is the change of log odds in the logistic regression. Probability of the dependent variable can be calculated from the odds ratio. When the exponential beta (odds ratio) is greater than one indicates that independent variable is an important risk factor and values close to zero also indicate important risk factors for the variable but it has negative effect on the dependent variable.

The selection of independent variables in the logistic regression model is also an important issue. For the selection of model variables, each of the independent variables is analyzed with a dependent variable using univariate logistic regression analysis. Then the variables that have probability values under 0.25 are proposed in the model. On the other hand, before analysis, the presence of multicollinearity among independent variables should be examined. If the correlation values among the variables are less than 0.90, it is assumed that there is no multicollinearity (Tabachnick and Fidell, 2007).

In this study, sex of calf, parity and calving year-season effects on stillbirth were analyzed with binary logistic regression analysis by the following model:

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)}}$$

where P(Y) is the probability of stillbirth levels, X_1, X_2 and X_3 sex of calf, parity and calving year-season effects, respectively, β_0, constant; β_1, β_2 and β_3 are the coefficients of regression. The first level of the independent variables was selected as reference category and compared to other levels. Odds ratios were used as a comparison criterion. The statistical significance of each coefficient (β) in the model was tested with the Wald-statistic has a chi-square distribution with degree of freedom equal to total number of covariates in the model to test the null hypothesis (Field, 2005):

$$W = \frac{\hat{\beta}}{S(\hat{\beta})}$$

where $\hat{\beta}$ represents the estimated parameters and $S(\hat{\beta})$ are their respective standard errors. The goodness-of-fit of the logistic regression model was evaluated with Hosmer and Lemeshow test statistic $(\hat{C}_g^*)$. The Hosmer-Lemeshow statistic evaluates the goodness-of-fit by creating 10 ordered groups of subjects and then compares the actual number in each group (observed) to the number predicted by the logistic regression model (predicted) and calculated as follows (Hosmer and Lemeshow, 2000):

$$\hat{C}_g^* = \sum_{k=0}^{1} \sum_{k=1}^{10} \frac{(O_{k_1} - E_{k_1})^2}{E_{k_1}}$$

The test statistic follows a chi-squared distribution with n-2 degrees of freedom. For the test statistic, non-significance outcome is desirable, which indicates that the model prediction does not significantly differ from the observed.

Binary logistic regression procedure in SPSS 20 was applied for descriptive statistics and for the factors affecting stillbirth, which were presented in this study, before the univariate logistic regression analysis was carried out for determining the potential risks of variables. Later the multivariate logistic analysis for the variables, which had P-value ≤0.25, was applied with the enter method that enters all variables at the same time.

RESULTS and DISCUSSION

The incidences of stillbirth for Holstein calves by different levels of sex, parity, calving year and year-season variables were shown in Table 1. The overall stillbirth rates for male and female calves were 51.4% and 48.6%, respectively. Although the stillbirth rate was similar for male and female calves, it was higher at first lactation than other lactations. Approximately half of the stillbirths (52.1%) were in the first lactation. On the other hand, the stillbirth rate of calves born from May to September 2004 was 41.8% and greater than other year-seasons (Table 1).

In this study, before applying binary logistic regression analysis, presence of multicollinearity among variables was examined. The Spearman correlation coefficients and their statistical significance are given in Table 2. As can be seen in Table 2, the correlations among variables were less than 0.70 (Tabachnick and Fidell, 2007) and accordingly it was determined that there was no multicollinearity among variables.

Table 1. Live and stillbirth frequencies of studied characters

Sex	Calf live frequencies	(%)	Stillbirth frequencies	(%)
Male	193520	51.4	14170	50.3
Female	182774	48.6	13995	49.7
Parity	-	-	-	-
1	111160	29.5	14670	52.1
2	114404	30.4	5637	20
3	72895	19.4	3492	12.4
4	77835	20.7	4366	15.5
5	21057	5.6	1207	4.3
6≤	15441	4.1	953	3.4
Year-Season	-	-	-	-
1	40446	10.7	3351	11.9
2	107150	28.5	7046	25
3	147933	39.3	11760	41.8
4	6402	1.7	421	1.5
5	74363	19.8	5587	19.8

Table 2. Spearman correlation coefficients and probabilities among sex of calf, parity and year-season of calving variables

		Parity	Year-Season
Sex	Correlation Coefficient	-0.02	-0.00
	Probability	0.00	0.69
Parity	Correlation Coefficient	1	-0.02
	Probability	-	0.00

Results of the univariate logistic regression analysis for the potential risk factors on stillbirth are shown in Table 3. As seen in Table 3, the P-values of the variable effects were found to be ≤0.25. Therefore, all variables were assumed to be risk factors and they were included in the multivariate logistic regression model. The results of the multivariate logistic regression analysis are shown in Table 4.

Compared to the reference category (male calves), the stillbirth rate of female calves was 1.05 times higher than male calves. Moreover, the risk of stillbirth decreased with increasing parity numbers. In addition, the stillbirth risk for calves born at third and fifth levels of year-season variable were found to be greater than for those born at second and fourth levels (Table 4).

Table 3. Univariate logistic regression results between stillbirth and sex of calf, parity and year-season of calving factors

Variables	B	SE	Wald	df	p	OR-Exp(B)	95% Confidence	
Female	0.05	0.01	13.09	1	0.00	1.05	1.02	1.07
Constant	-2.61	0.01	90235.12	1	0.00	0.07		
Parity			5896.28	5	0.00			
2	-0.99	0.02	3686.46	1	0.00	0.37	0.36	0.39
3	-1.01	0.02	2722.17	1	0.00	0.36	0.35	0.38
4	-0.91	0.02	1477.96	1	0.00	0.40	0.39	0.42
5	-0.83	0.03	729.62	1	0.00	0.43	0.41	0.46
≤6	-0.76	0.04	484.87	1	0.00	0.47	0.44	0.50
Constant	-2.03	0.01	53151.52	1	0.00	0.13		
Year-Season			188.73	4	0.00			
2	-0.23	0.02	112.54	1	0.00	0.79	0.76	0.83
3	-0.04	0.02	4.12	1	0.04	0.96	0.92	1.00
4	-0.23	0.05	18.70	1	0.00	0.79	0.72	0.88
5	-0.10	0.02	18.55	1	0.00	0.91	0.87	0.95
Constant	-2.49	0.02	19197.81	1	0.00	0.08		

Table 4. Binary logistic regression results for stillbirth rate of calves

Variables	N	B	SE	Wald	df	p	OR-Exp(B)	95% Confidence Intervals	
Male	207690	-	-	-	-	-	Reference	-	-
Female	196770	0.03	0.01	4.38	1	0.04	1.03	1.00	1.05
Parity	404460			5885.85	5	0.00		-	-
1	125830	-	-	-	-	-	Reference	-	-
2	120041	-0.98	0.02	3675.80	1	0.00	0.37	0.36	0.39
3	76387	-1.01	0.02	2714.39	1	0.00	0.36	0.35	0.38
4	43543	-0.91	0.02	1482.47	1	0.00	0.40	0.39	0.42
5	22264	-0.84	0.03	734.43	1	0.00	0.43	0.41	0.46
≤6	16395	-0.76	0.04	486.72	1	0.00	0.47	0.44	0.50
Year-Season				187.41	4	0.00		-	-
1	43797	-	-	-	-	-	Reference	-	-
2	114196	-0.24	0.02	115.41	1	0.00	0.79	0.76	0.83
3	159693	-0.05	0.02	6.63	1	0.01	0.95	0.91	0.99
4	6823	-0.29	0.05	28.70	1	0.00	0.75	0.67	0.83
5	79951	-0.11	0.02	24.06	1	0.00	0.89	0.86	0.94
Constant	-	-1.93	0.02	9074.52	1	0.00	0.15	-	-

Hosmer-Lemeshow test observed and expected values of the decimal risk group are presented in Table 5. Hosmer-Lemeshow statistics related with these risk groups are calculated as follows:

$$\hat{C}_g^* = \frac{(41301 - 41322.140)^2}{41322.140} + \frac{(35763 - 35723.236)^2}{35723.236} + ... + \frac{(7886 - 7844.191)^2}{7844.191} = 12.68$$

Table 5. Contingency table for Hosmer and Lemeshow test

	Calf live		Stillbirth		Total
	Observed	Expected	Observed	Expected	Observed
1	41301	41322.140	1777	1755.860	43078
2	35763	35723.236	1595	1634.764	37358
3	33127	33157.704	1661	1630.296	34788
4	31407	31424.089	1615	1597.911	33022
5	31385	31348.515	1591	1627.485	32976
6	39273	39318.071	2152	2106.929	41425
7	36990	36959.199	2091	2121.801	39081
8	33542	33636.408	3147	3052.592	36689
9	38127	37983.829	4650	4793.171	42777
10	55379	55420.809	7886	7844.191	63265

The results of Hosmer-Lemeshow test statistic and its probability are given in Table 6. From these results it can be said that the logistic regression model provided a good fit.

Table 6. Chi-Square value and probability of Hosmer-Lemeshow test statistic

Step	Chi-Square value	df	p
1	12.68	8	0.12

In this study sex, parity and calving year - season variables are examined to determine if these are risk factors or not on stillbirth ratios. The effects of variables were determined by logistic regression analysis in which all factor were contained. According to the results of analysis, all levels of sex, parity and calving year-season variables were found to have a significant effect on stillbirth rates. Similar results have been reported by Meyer et al. (2001), Bicalho et al. (2007) and Atashi (2011).

When the variables were examined in detail, stillbirth rates were found to be similar for male and female calves. However, by fitting the logistic regression model, females have 3% higher stillbirth rates than males. This result is not similar to that of Bicalho et al. (2001), who reported that stillbirth rates of female calves were 23% lower than those of male calves. Meyer et al. (2001) reported that stillbirth rates of female calves were 7% lower than male calves for primiparous cows. However, female calves of multiparous cows have a stillbirth rate 12% higher than that of male calves. These results are higher than our results, but show a similar trend. On the other hand, in this study the effect of parity was found to be an important risk for stillbirth, but the risk was decreased by increased number of parity. Indeed, stillbirth rates of 2nd, 3rd, 4th, 5th and 6th lactations were significantly lower than first lactation ($P<0.05$).

Compared to the first lactation (as a reference category) the stillbirth risks of other lactations were 0.37, 0.36, 0.40, 0.43 and 0.47 times lower, respectively. These results are consistent with the results of Meyer et al. (2001), Bicalho et al. (2007) and Atashi (2011).

This study showed that the year-season effect was associated with the risk of stillbirth. Compared to stillbirths in the months of May to September 2003 (as a reference category), the other levels of year-season effects showed 0.79, 0.95, 0.75 and 0.89 times lower stillbirth rates, respectively. Meyer et al. (2001) reported that stillbirth risk was higher in summer than in the winter season. This result is in agreement with our findings that the summer season had higher risk of stillbirth. However, Atashi (2011) found the calving season did not significantly affect stillbirths.

As a result, calving mortality can be reduced by improvements in herd management considering the mentioned risk factors. However, there are other factors that may affect stillbirth such as pelvic size of dam, gestation length, birth weight of calf and nutrition etc. Knowledge of these factor effects is important to keep stillbirth under control. On the other hand, there are not adequate and reliable statistics on calf death and its causes. The stillbirth ratio ranged between from 2% to 48% in the EU countries. In the United States, almost 7% of Holstein calves died within the first 48 hours of birth (Meyer, Berger and Koehler, 2000). In addition, it was found that the birth of dead calves caused $125 million/year in economic loss (Salfer, 2005). Much consideration should be given to birth records for accurate decisions. Reducing stillbirths by breeding programs will increase productivity and at the same time benefit the farmer economy.

ACKNOWLEDGEMENTS

The authors wish to thank the US National Association of Animal Breeders for providing data.

REFERENCES

Antonogeorgos G, D.B. Panagiotakos, K.N. Priftis and A. Tzonou. 2009. Logistic regression and linear discriminant analyses in evaluating factors associated with asthma prevalence among 10- to 12-years-old children: divergence and similarity of the two statistical methods. International Journal of Pediatrics, 2009:2009.

Atashi H. 2011. Factors affecting stillbirth and effects of stillbirth on subsequent lactation performance in a Holstein dairy herd in İsfahan. Iranian Journal of Veterinary Research, Shiraz University 12:24-30.

Bicalho RC, K.N. Galvao, S.H. Cheong, R.O. Gilbert, L.D. Warnick and C.L. Guard. 2007. Effect of stillbirth on dam survival and reproduction performance in Holstein dairy cows. Journal of Dairy Science. 90:2797-2803.

Çokluk Ö, G. Şekercioğlu and S. Büyüköztürk. 2010. Multivariable Statistics for Social Sciences (In Turkish). Pegem Akademi Yayıncılık, Ankara.

Cook D, P. W.M. Dixon, M.S. Duckworth, K. Kaiser, W.Q. Koehler, W.Q. Meeker and W.R. Stephenson. 2001. Binary response and logistic regression analysis. Part of the Iowa State University NSF/ILI project, Beyond Traditional Statistical Methods. Available at http://www.public.iastate. edu/~stat415/ stephenson/stat415_chapter3.pdf (accessed 3 August 2015).

Field A. 2005. Discovering Statistics Using SPSS. 2nd ed. London: Sage.

Hair J, B. B. Black, B. Babin, R. Anderson and R. Tatham. 2006. Multivariate Data Analysis, 6th ed. Upper Saddle River, NJ: Prentice-Hall.

Hosmer D.W. and S. Lemeshow. 2000. Applied Logistic Regression. John Wiley & Sons, New York.

Johnson R.A. and D.W. Wichern. 2005. Applied Multivariate Statistical Analysis, Prentice Hall, Upper Saddle River, New Jersey.

Mertler C.A. and R.A. Vannatta. 2005. Advanced and Multivariate Statistical Methods: Practical Application and Interpretation, 3rd ed. Pyrczak, Los Angeles.

Meyer C.L., P.J. Berger, K.J. Koehler , J.R. Thompson and C.G. Sattler. 2001. Phenotypic trends in incidence of stillbirth for Holsteins in the United States. Journal of Dairy Science. 84:515-523.

Meyer C.L., P.J. Berger and K.J. Koehler. 2000. Interactions among factors affecting stillbirths in Holstein cattle in the United States. Journal of Dairy Science. 83:2657-2663.

Park, H.M. 2010. Regression models for binary dependent variables using STATA, SAS, R, LIMDEP, and SPSS. The university information technology services (UITS) center for statistical and mathematical computing, Indiana University. Available at http://www.iu.edu/~statmath/stat/all/cdvm/cdvm.pdf (accessed 3 August 2015).

Salfer J. 2005. "Don't kill your calves with kindness" *Dairy Star*, Available at http://www.extension.umn.edu/agriculture/ dairy/health-and-comfort/dont-kill-your-calves-with-kindness/ (accessed 3 August 2015).

Sharma S. 1996. Applied Multivariate Techniques, John Wiley & Sons, USA.

Stephenson B. 2008. Binary response and logistic regression analysis. Available at http://www.public.iastate.edu/ ~stat415/Stephenson/ stat415_chapter3.pdf (accessed 3 August 2015).

Tabachnick B.G. and Fidell L.S. 2007. Using Multivariate Statistics, 5th ed. Pearson.

Yakubu A., M.M. Muhammed and I.S. Musa-Azara. 2014. Application of multivariate logistic regression model to assess factors of importance influencing prevalence of abortion and stillbirth in Nigerian Goat Breeds. Biotechnology Animal Husbandry 30:79-88.

Zadeh N.G.H. 2014. Effect of dystocia on the productive performance and calf stillbirth in Iranian Holsteins. Journal of Agricultural Science and Technology 16:69-78.